About

This document will give a short insight about the project we have been given and how we plan to realise the outcome of the project.

Building blocks

Group 3



Wensley Podrono Heather van Diemen Donovan V.D Khoun Eelke Sneijders

# The #1 in Consumer AI - Building BlocksPreface

We are BB Royalty from Fontys, a group of Data Science enthusiasts. Building Blocks had approached Fontys with the question if people were interested to solve a data science related problem. One of their clients has had problems with creating profiles that were applicable to the customers of their webshop. The reason the profiles were not succesful was that they were too generic. Building Blocks then posed the question to us, the students, if we could find a way to create more accurate profiles. Hence, why we came together to find a solution. Building Blocks provides the data from the client to us and we will look into the usable patterns.

Contents

[1 Preface 1](#_Toc101433185)

[3 Phase 1: Proposal 4](#_Toc101433186)

[3.1 Stakeholders and domain understanding 4](#_Toc101433187)

[3.2 Possible Restrictions 4](#_Toc101433188)

[3.2.1 What limits do we have in our project? 4](#_Toc101433189)

[3.3 Data Sourcing 4](#_Toc101433190)

[3.4 Analytic Approach 5](#_Toc101433191)

[4 Phase 2: Provisioning 6](#_Toc101433192)

[4.1 Data Requirements 6](#_Toc101433193)

[4.1.1 Proposal, Domain Understanding 6](#_Toc101433194)

[4.1.2 Find the Stakeholders 7](#_Toc101433195)

[4.1.3 Identify required data elements 7](#_Toc101433196)

[4.1.4 Identify candidate data sources 7](#_Toc101433197)

[4.2 Data Collection 8](#_Toc101433198)

[4.2.1 Determine what information you want to collect 8](#_Toc101433199)

[4.2.2 Where to store your retrieved data? 8](#_Toc101433200)

[4.2.3 Use a (traceable) version and naming system 8](#_Toc101433201)

[4.2.4 Determine how and how often you want to retrieve the data 8](#_Toc101433202)

[Appendix A: Stakeholders 9](#_Toc101433203)

[5 Building blocks 9](#_Toc101433204)

[5.1 What is their mission? 9](#_Toc101433205)

[5.1.1 How does our project relate to BB? 9](#_Toc101433206)

[5.1.2 Who are their clients? 9](#_Toc101433207)

[5.1.3 What do they expect to be the result of the project? 9](#_Toc101433208)

[6 Otto 10](#_Toc101433209)

[6.1 What is Otto? 10](#_Toc101433210)

[6.1.1 Wat is a mail order business? 10](#_Toc101433211)

[6.1.2 How does our project relate to Otto? 10](#_Toc101433212)

[6.1.3 Who are their clients? 10](#_Toc101433213)

[6.1.4 Why does Otto want personalization? 10](#_Toc101433214)

[Appendix B: Domain Understanding 11](#_Toc101433215)

[7 Domain Understanding 11](#_Toc101433216)

[7.1 E-commerce 11](#_Toc101433217)

[7.2 Personalization 11](#_Toc101433218)

[7.2.1 Why utilize personalization as a retailer? 11](#_Toc101433219)

[7.3 Customer experience 11](#_Toc101433220)

[7.3.1 What are some common examples of improving customer experience through personalization? 11](#_Toc101433221)

[7.3.2 What kind of recommendation do customers want? 11](#_Toc101433222)

[7.4 Personalized e-commerce 12](#_Toc101433223)

[7.4.1 What are the most common types of personalization in ecommerce? 12](#_Toc101433224)

[7.5 Personalized marketing campaigns 12](#_Toc101433225)

[7.6 How does it work? 12](#_Toc101433226)

[7.7 Benefits of personalized marketing 13](#_Toc101433227)

[8 Sources 14](#_Toc101433228)

# Phase 1: Proposal

The proposal covers 4 components:

* Domain Understanding (can be found in appendix A and B)
* Possible restrictions
* Data Sourcing
* Analytic Approach

## Stakeholders and domain understanding

*To improve the daily practices of domain experts, you need to know as much as possible about the specific domain. Thus, a thorough exploration and understanding of the domain in which you work, gained through different types of research, is always the starting point of any AI project.*

Building Blocks helps other companies to better serve their customers through AI. In this case Building Blocks will help Otto, an online retailer. For more information about the stakeholders, see Appendix A.

The goal of this project is to find personalization opportunities in the data and present information about groups of customers and/or individual customers which Otto can use for personalization. Further information about the domain understanding can be found in Appendix B.

## Possible Restrictions

### What limits do we have in our project?

In this project we have some restrictions that we must consider before we start working on it. Here are a few we know of:

* The data we have can only be stored locally on a VM.
* It is hard to work in one notebook with multiple people if data is stored locally.
* For now, we only have one day a week to work on the project (This will increase later).
* The VM we use has low CPU power which makes it impossible to work in with multiple people at the same time.
* We have received a dataset from Otto, this is the only dataset we can use and there will be no further data provided.

## Data Sourcing

*This section describes what kind of data we have and how we got this data.*

Because we are doing this assignment for the company ‘building blocks’, we receive the data we are going to use from them. The dataset is about how often people click on a certain item in a web shop. In this case that web shop is Otto. Based on this data we can do a trend analysis after which we can target certain groups more specifically.

The dataset is found on a virtual machine, which Building Blocks provide us with. On this virtual machine, not more than two people can work on the project at the same time. The dataset cannot be taken off the virtual machine for any reason, the data is confidential and cannot be leaked.

## Analytic Approach

*The next step in writing a proposal is to define a specific goal and approach for the project.*

Otto is a wholesale company. They want to increase their sales by using personalized profiles for each (registered) customer. By making purchases or adding items to your wish list, Otto is enabled to create these profiles. These, if being handled correctly, can be used to recommend items in the catalogue to your customer. The problem however is that these are too generic: customers are not inclined to buy the recommendations.

If the profiles you have created are being too generic and you try to recommend something based on them, you will not be successful. Why is that?

Empathy will scale. Empathy is the ability to relate to and understand another person’s emotions. It is the basis of strong relationships. Understanding social cues and adapting to them is how people build trust. That is not easy to do digitally or at scale.

Forming relations with your customers will result in a higher revenue due to the practice of personalized advertisement. When we feel understood, we are more inclined to do something. It invokes a positive feeling which will then be associated with the company. In return, customers will buy more.

The smart thing to do would be to see how the data is organized, what pre-existing patterns exist within the dataset. Use those to our advantage and then label them and see how we can predict future patterns. There must be target indicators which are the main factors on which we can base our profiles.

The personalization is going to play a key factor in the success of our project. If we can spot patterns that could represent the likeliness of people buying a product, then we will be successful. We will make use of two types of profiles, a profile for what type of user we have and a profile for the type of product. The indicators we are going to utilize are: categories, websiteEventids, the price and the amount it has been clicked. We assume that the combination of these four indicators could reveal a pattern that correlates with a profile.

Another factor we could consider is the time. The possibility of correlation between the type of product or user and time is present. Our assumption is that some users tend to buy more during certain times/dates and some products are more successful during a time and date. That would mean that we could also look into what holiday it is or what weather it is.

By not limiting ourselves to just the users or just the products, different types of patterns could become visible. By having different perspectives, we can make an assessment which is more representative of the real world.

# Phase 2: Provisioning

In phase 2: Provisioning, the following four components occur:

1. Data Requirements
2. Data Collection
3. Data Understanding
4. Data Preparation

## Data Requirements

The essence of Data Requirements is to understand that a stakeholder is storing this data for a specific reason or purpose

### Proposal, Domain Understanding

*Make a list of the domains and describe the relation between them*

The domain we need to collect data from is Otto B.V.

The data we received contains data about the products of the web shop. This dataset contains the following information about the product:

* Product Id
* Product Group
* Brand
* Price
* Assortments
* Product Categories
* Image URL

The other data we received from Otto B.V. contains information about the web events happening at the Otto web shop. This dataset contains the following information about the web events:

* Event Id
* Website Visitor Id
* Website Visit Id
* Hit Number
* Visit Start Datetime
* Event Datetime
* Event type
* Product Id
* Base Product Id
* Search Category
* Category View Category

By linking both of the datasets together, we can easily see a pattern of items the customers click on. We can link both datasets by product id. This way we can see what categories customers are mostly searching for and what kind of product. By examining what certain customers search for and clicks on, we can also estimate the target audience (male, female, child).

### Find the Stakeholders

*From whom is the storage of data beneficial in the domain you are researching? Try to approach this creatively. Don’t look only inside your research domains but also in adjacent ones. There is a possibility you’ll find something unexpected.*

**Stakeholders:**

* Building Blocks
* Fontys Hogeschool
* Otto B.V.

The storage of de data is beneficial for Otto B.V. because they have this data stored, they can target advertising to the right audience and the right season. In this way, they will sell more products considering consumers are more likely to buy recommended products if it suits them. More about stakeholders can be found the Appendix A.

### Identify required data elements

*Figure out the facts and dimensions you require to answer your research. Multiple facts can only be joined through shared dimensions, there are specific procedures you must obey to make these joins work!*

To make proper categorizations, having a clear view of what the categories are, is essential. Otherwise, correlations will not show up because the algorithm will not be able to acknowledge the fact that two items might share categories.

Knowing what the event on the website is, is also important. A bought item is more important than an item one of the customers scrolled through.

To figure out a profile for the customer, having a log of time is nice: you can figure out if they are night shoppers, holiday haulers or just frequent micro shoppers. In recognizing these patterns, we can create more appropriate profiles.

### Identify candidate data sources

*Combine all the facts and then document each potential data source*

Building Blocks/Otto are the suppliers of the current dataset we are working with. Otto is the origin of our data: a dataset regarding the sales of their items. If we were to look for more datasets which we would use to tweak our model, it would be a company meeting the data requirements. A wholesale company would not have to be a necessity. We could also combine the dataset with the sales during for example the holidays. The sales in a specific region. Personal interests when having taken an inquiry are also something we could make use of to tweak the user profiles.

## Data Collection

Data collection is the process of gathering information from an established system. The goal for all data collection is to capture quality evidence that allows analysis to lead to the formulation of convincing and credible answers to the questions that have been posted.

### Determine what information you want to collect

*Set a time frame you want to look for. Try to be as specific as you can. This is a premise you can use for looking at the right location*

Building Blocks provided us with a dataset that contains data of all the web events that were happening at the Otto web shop starting from date: 2021-11-23 to date 2022-01-03. We don’t need a longer time frame for this project. The dataset that is about the products does not use a time frame.

### Where to store your retrieved data?

*Where are we going to store out data. Think about 24/7 availability of the data if you work in a group.*

As mentioned in Phase 1, the data we need for this project is on a virtual machine that is provided by building blocks. We can access this machine individually at any time to work with the data.

### Use a (traceable) version and naming system

*Because you need to refer to the dataset frequently, you must think of an easy identifiable naming system.*

For this project we named the dataset that is about the products: ‘productFrame’. We named the dataset that is about the web events: ‘eventFrame’. Other names we will use for this project will be explained in the notebook we are using. After each and every iteration we will export our jupyter notebook in a html/pdf format with date and time. This way we can keep track of out versions.

### Determine how and how often you want to retrieve the data

*This depends highly on the quantity of data or load and update strategy of the source.*

Otto gave us a dataset with the dates of the month of December, the last days of the month of October and the first days of the month of January. This allows us to make a prediction about the month of December in the year 2022. When we can get more data from Otto about the other months, we can also make predictions for the rest of the year. As written in the proposal, this is not the assignment for this project so we will stick to these months and will not receive any new data.

1. Stakeholders

# Building blocks

## What is their mission?

*“The #1 in Consumer AI (Artificial Intelligence). We empower companies to act in consumers’ favor.” (z.d.). Building blocks.* Building blocks seeks to better consumer experience by helping their customers with AI. This, according to building blocks' website, helps grow businesses, enhance brand experiences, and create loyal customers. They want to create a system that makes it a solution where everyone benefits.

### How does our project relate to BB?

We are going to work on a project much like Building Blocks would themselves for Otto. The plan is to improve Otto’s customer’s experience, so they are more likely to buy products related to their user profile. This will help hold customers and to have them buy more products they otherwise might not have found.

### Who are their clients?

Their clients are companies who want to increase customer experience using AI on website or other aspects. They focus on commercial companies that are seeking to better their brand and to increase revenue. These companies can vary wildly but some examples are Basic fit, Corendon, Coca Cola and Samsung.

### What do they expect to be the result of the project?

Information about Otto customers to achieve personalization when interacting with Otto. Personalization could be recommending items in-store based on individual preferences. Information could be explicit customer preferences about products. It could mean patterns at what hour the customer is visiting the website.

# Otto

## What is Otto?

Otto B.V. is a Dutch online store. The company started as a mail order company, and it is part of the German Otto group. Otto B.V. is in Tilburg. The company was founded in 1979 and was one of the first mail order companies in the Netherlands. The founder is Michael Otto. In the fall of 2010 Otto stopped publishing the traditional catalogue and went completely to online operations.

### Wat is a mail order business?

Mail order is the remote purchase of goods that are then delivered by mail or courier. Orders can be placed on a form, by mail, fax, e-mail, or telephone. Nowadays orders are increasingly made via the Internet, the products are then delivered to your home.

### How does our project relate to Otto?

Otto wants to improve its customer experience by personalizing the interaction a costumer has with Otto through AI.

### Who are their clients?

Otto is focused on fashion, multimedia, and living items. Anyone who’s interested in those categories is a potential customer. To which demographic their actual customers belong to is only known to Otto. Otto has not shared demographic data with us. Their potential and actual customers are most likely to be in the Netherlands only. Otto only ships to Dutch addresses, but also mentions how to order from outside the country.

### Why does Otto want personalization?

Since Building Blocks wants to change the way consumers are treated for the better, we can safely assume Otto want to utilize personalization for the sake of the consumer. Applying personalization to improve customer experience is both in the interest of the consumer and in the interest of Otto. Offering a highly personalized experience to the customers will help to differentiate Otto from other online retailers and will help Otto to gain a competitive advantage.

1. Domain Understanding

# Domain Understanding

## E-commerce

E-commerce is buying and selling of products and services over the internet.[1] For this project Otto is only active in online shopping. This is an area of e-commerce in which consumers buy from a seller directly on the internet using a web browser or an application.[2] Devices that are used are desktop computers, laptops, tablets and smartphones.[2] Consumers can find products by visiting the website of Otto directly.[2] Consumers can also find a product that Otto is selling by searching for a product through a shopping search engine.[2]

## Personalization

Salesforce gives the following definition of personalization: “Personalization is the act of tailoring an experience or communication based on information a company has learned about an individual.”[3]

### Why utilize personalization as a retailer?

Marketers believe personalization is important for customer relation.[4] Customers may also expect experiences tailored to them.[4] The main reason for utilizing personalization is for improving the customer experience.[4] It pays to treating the customer well.[4] There are other reasons for a retailer to utilize personalization such as increasing return on investment and converting potential customers to actual customers.

## Customer experience

### What are some common examples of improving customer experience through personalization?

Personalization used to be only employed when advertising targeted offers.[5] Personalization is now present in the customer experience as a whole.[5] Amazon shows customers products that are often bought with the items the customer is viewing.[5] This is different from recommending items that are bought by other users when browsing through items. The former is a recommendation more based on individual factor, whereas the latter is less geared towards the individual

### What kind of recommendation do customers want?

Customers want to receive recommendations that are relevant to them individually.[6] This is opposed to receiving recommendations based on similar customers.[6]

## Personalized e-commerce

### What are the most common types of personalization in ecommerce?

Ecommerce personalization is all about creating a more tailored experience for your customer. It’s something that you can do across the entire user journey to make it relevant to each person that visits your site. The right eCommerce personalization strategy helps you define your traffic’s behavior during their shopping session as precisely as possible.

This goes hand in hand with Personalized Ads. It all boils down to making use of data you have collected and then personalizing the content for the visitor/prospect customer.

Examples of common types:

1. Create Offers Based on User Behavior
2. Display Products or Campaigns Based on Location
3. Upsell, Cross-sell, & Downsell at (& after) Checkout
4. Help Users Navigate Your Website
5. Personalize Your Opt-in Campaigns & Emails

## Personalized marketing campaigns

In the modern world, we can make use of personalizing advertisement. To keep it brief: it is more than just inserting the name of the customer in a generic email, it is about reaching the right audience at the right moment with the right suggestions.

## How does it work?

The way you make it work is by tailoring the marketing based on the data you've collected from your contacts or your own dataset. This includes interests, shopping preference, purchase history etc. Companies often have a team specialized for the personalized marketing. (Experts in the field, people who know tech, strategic campaign designers who have experience with personalization).

As we mentioned, we need data to tailor the marketing. They either cultivate it themselves or purchase the data from a business partner. You must make sure you handle the data with care. You do not want to create a rift between you and your (potential) customers.

Segment your audience using the data. Do this based on the patterns that arise in the data you have acquired. This way, you are enabled to reach the right audience.

Examples of personalized marketing campaigns:

* Custom emails: By utilizing personalized ads, you can simulate a feeling of empathy which in return can drive your revenue.
* Targeted discounts: \*According to a 2017 report, over 60% of consumers say getting a discount within an hour of interacting with a brand can help drive loyalty.
* Less generic advertising: \*With the rise of personalization, consumers are becoming increasingly less accepting of generic ads: Impersonal shopping experiences frustrate three-quarters of customers. On the other hand, over 70% of shoppers respond to marketing only when it is customized to their interests.

## Benefits of personalized marketing

What benefits from personalized marketing?

It will offer a more satisfying customer experience: the more we feel that someone understands our needs, the more inclined we are to trust them with our problems.

The customers’ purchase(s) can be used for other websites to personalize it even further. This allows you to sell the acquired data. Vice versa, whenever you can acquire data, you are able to tailor your marketing to the needs of the customers thus increase your revenue.

As we had mentioned, your marketing can be tailored. A side effect of this is also that the customers you have already garnered will feel more valued by the company. Their needs are being understood, so a sort of bond has come to fruition. By feeling this way, you can make the customers come back more often and turn it into more revenue and collect more data.

It helps customers discover new products: some prospective customers do not even know they need a specific type of product, but their browsing behavior could resemble that of someone else who has already bought their “needed” item. Making use of the patterns that exist among humans, we can suggest products to a lot of people.

Another benefit for the company is that it is more cost effective than traditional marketing. Traditional marketing is intended for a broad range of people whilst personalized marketing plays into the desires of a specific group. This way, you can invest fewer resources into creating success which is also repeatable.

# Sources

*Bedrijfsprofiel | OTTO*. (z.d.). Geraadpleegd op 25 februari 2022, van https://www.otto.nl/special/bedrijfsprofiel/

Clark, S. (2020, 17 augustus). *6 Personalization Tips for a Better Customer Experience*. Geraadpleegd op 7 maart 2022, van https://www.cmswire.com/customer-experience/6-personalization-tips-for-a-better-customer-experience/

*Do people really want personalised ads online?* (z.d.). Geraadpleegd op 7 maart 2022, van https://www.globalwitness.org/en/blog/do-people-really-want-personalised-ads-online/

Lee, L. (2021, 31 mei). *Every Move You Make. . . I’ll Be Watching You: How Personalized Ads Work*. Geraadpleegd op 7 maart 2022, van https://enveritasgroup.com/campfire/why-am-i-seeing-this-ad-how-personalized-ads-work/

Lindecrantz, E., Pian Gi, M. T., & Zerbi, S. (2020, 28 april). *Personalizing the customer experience: Driving differentiation in retail*. Geraadpleegd op 7 maart 2022, van https://www.mckinsey.com/industries/retail/our-insights/personalizing-the-customer-experience-driving-differentiation-in-retail

*Otto BV Wikipedia*. (z.d.). Geraadpleegd op 25 februari 2022, van https://nl.wikipedia.org/wiki/Otto\_B.V.

*Personalization Defined: What is Personalization?* (z.d.). Geraadpleegd op 7 maart 2022, van https://www.salesforce.com/resources/articles/personalization-definition/

*Personalized advertising - Advertising Policies Help*. (z.d.). Geraadpleegd op 7 maart 2022, van https://support.google.com/adspolicy/answer/143465?hl=en

*The #1 in Consumer AI*. (2022, 16 maart). Geraadpleegd op 7 maart 2022, van https://building-blocks.com/

*What is Personalized Marketing?* (z.d.). Geraadpleegd op 7 maart 2022, van https://mailchimp.com/marketing-glossary/personalized-marketing/

Wikipedia contributors. (2022a, 16 januari). *Online shopping*. Geraadpleegd op 7 maart 2022, van https://en.wikipedia.org/w/index.php?title=Online\_shopping&oldid=1065935472

Wikipedia contributors. (2022b, 12 februari). *E-commerce*. Geraadpleegd op 7 maart 2022, van https://en.wikipedia.org/w/index.php?title=E-commerce&oldid=1071315480

Wikipedia-bijdragers. (2020a, 30 mei). *Postorder*. Geraadpleegd op 25 februari 2022, van https://nl.wikipedia.org/wiki/Postorder

Wikipedia-bijdragers. (2020b, 18 december). *Otto Group*. Geraadpleegd op 25 februari 2022, van https://nl.wikipedia.org/wiki/Otto\_Group

Zimmerman, A. (2020, 29 mei). *Why the Time Is Right for Personalization*. Geraadpleegd op 7 maart 2022, van <https://www.cmswire.com/digital-marketing/why-the-time-is-right-for-personalization/>

https://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/the-future-of-personalization-and-how-to-get-ready-for-it

Geraadpleegd op 24 maart 2022